```
OM protein - nucleic search, using frame_plus_p2n model

                                    Copyright (c) 1993 - 2006  Bioacceleration ltd.
                                    GenCore version 5.1.7

Run on:        February 28, 2006, 12:23:55 ; Search time 779.148 Seconds
                                                 (without alignments)
                                             3470.568 Million cell updates/sec

Title:           US-10-511-270-2
Perfect score:   1702
Sequence:        1 MLGPQVWSSVRQGLSRSLSR.........ELSPAEEALRMDFTSNGWL 327

Scoring table:   BLOSUM62
                 Xgapop  10.0 , Xgapext  0.5
                 Ygapop  10.0 , Ygapext  0.5
                 Fgapop   6.0 , Fgapext  7.0
                 Delop    6.0 , Delext   7.0

Searched:        9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters:    19587084

Minimum DB seq length:  0
Maximum DB seq length:  2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
-MODEL=frame+_p2n.model  -DEV=xlp
-Q=/abs8/ABSSWEB spool/US10511270/runat 27022006 164254 20192/app_query.fasta_1
-DB=Published Applications NA_Main  -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0  -UNITS=bits -START=1  -END=-1
-MATRIX=blosum62  -TRANS=human40.cdi  -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100  -THR_MIN=0  -ALIGN=15  -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0  -MAXLEN=2000000000  -HOST=abs805p
-USER=US10511270 @CGN_1_1_2064 @runat_27022006 164254 20192  -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0  -WAIT  -DSPBLOCK=100  -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10  -XGAPEXT=0.5  -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
Published Applications NA_Main:*
  1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
  2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
  3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
  4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
  5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
  6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
  7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
  8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
  9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1696	99.6	2032	8	US-10-128-558-103	Sequence 103, App
2	1659	97.5	1061	9	US-10-511-270-1	Sequence 1, Appli
3	1641.5	96.4	1521	9	US-10-120-988-156	Sequence 156, App
4	1497	88.0	1017	9	US-10-511-270-3	Sequence 3, Appli
5	1430	84.0	1336	7	US-10-433-802-21	Sequence 21, Appl
6	1402	82.4	1235	6	US-10-006-285-453	Sequence 453, App
7	1034	60.8	2581	6	US-10-108-260A-344	Sequence 344, App
8	727	42.7	664	7	US-10-424-599-136324	Sequence 136324,
9	720	42.3	456	8	US-10-128-558-317	Sequence 317, Appl
10	619	36.4	389	3	US-09-918-995-8256	Sequence 8256, App
11	605.5	35.6	54718	8	US-10-719-993-7038	Sequence 7038, Ap
12	570	33.5	867	8	US-10-006-285-277	Sequence 277, App
13	517.5	30.4	348	8	US-10-369-493-42239	Sequence 42239, A
14	382	22.4	942	6	US-10-369-493-33440	Sequence 33440, A
15	376.5	22.1	13207	8	US-10-369-493-40233	Sequence 40233, A
16	375	22.0	23349	8	US-10-719-993-7016	Sequence 7016, Ap
17	375	22.0	245	6	US-10-006-285-111	Sequence 111, App
18	336.5	19.8	870	6	US-10-369-493-45137	Sequence 45137, Ap
19	334.5	19.7	933	6	US-10-369-493-39120	Sequence 39120, A
20	331.5	19.5	897	6	US-10-369-493-35026	Sequence 35026, A
21	331	19.4	897	6	US-10-369-493-31995	Sequence 31995, A
22	329	19.3	865	6	US-10-369-493-37079	Sequence 37079, A
23	325.5	19.1	897	6	US-10-156-761-6320	Sequence 6320, Ap
24	325	19.1	865	6	US-10-156-761-1	Sequence 1, Appl
25	325	19.1	9025608	6	US-10-282-122A-39025	Sequence 39025, Ap
26	323.5	19.1	906	7	US-10-156-761-2503	Sequence 2503, Ap
27	322.5	18.9	897	6	US-10-369-493-36973	Sequence 36973, A
28	322.5	18.9	925	6	US-10-369-493-39488	Sequence 39488, A
29	321	18.9	816	6	US-10-369-493-24858	Sequence 24858, A
30	320.5	18.8	852	6	US-09-070-927A-322	Sequence 322, App
31	315.5	18.8	5069	3	US-09-070-927A-322	Sequence 1, Appli
32	315.5	18.5	873	6	US-10-214-556-1	Sequence 1373, Ap
33	313.5	18.4	873	6	US-11-073-741-1	Sequence 24687, A
34	313.5	18.4	870	6	US-10-369-493-24687	Sequence 33837, A
35	312	18.3	867	6	US-10-369-493-33837	Sequence 41002, A
36	311	18.3	879	6	US-10-369-493-41002	Sequence 1373, Ap
37	310.5	18.2	4737	6	US-10-369-493-31553	Sequence 31553, A
38	305	17.9	885	6	US-10-238-075-1373	Sequence 509, App
39	302	17.7	750	3	US-09-974-300-509	Sequence 40937, A
40	301.5	17.7	888	6	US-10-369-493-40937	Sequence 37575, A
41	301.5	17.7	873	6	US-10-369-493-37575	Sequence 67, Appl
42	300.5	17.6	9170	9	US-10-650-274-67	Sequence 9404, Ap
43	299	17.5	894	6	US-10-282-122A-9404	Sequence 1374, Ap
44	298.5	17.5	882	6	US-10-238-075-1374	Sequence 8, Appli
45	293.5	17.2	495269	7	US-10-398-221-8	

ALIGNMENTS

RESULT 1
US-10-128-558-103
Sequence 103, Application US/10128558
Publication No. US20040219521A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Wang, Zhiwei
APPLICANT: Weng, Gezhi
APPLICANT: Boyle, Bryan J
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 812A
CURRENT APPLICATION NUMBER: US/10/128,558
CURRENT FILING DATE: 2002-04-22
PRIOR APPLICATION NUMBER: US/10/128,558
PRIOR FILING DATE: 2002-04-25
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: PCT/US00/35017
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 412
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 103
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (419)..(1402)
US-10-128-558-103

US-10-511-270-2 (1-327) x US-10-128-558-103 (1-2032)

Alignment Scores:

Pred. No.:	2.17e-193	
Score:	1696.00	Length: 2032
Percent Similarity:	99.7%	Matches: 326
Best Local Similarity:	99.7%	Conservative: 0
Query Match:	99.6%	Mismatches: 0
DB:		Indels: 1
		Gaps: 0

```
Qy     1 MetLeuGlyProGlnValTrpSerSerValArgGlnGlyLeuSerArgSerLeuSerArg  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 ATGCTGGGTCCCCAAGTCTGGTCTTCGTGTAGGCAGGGGGCTCAAGCGGAGCTTGTCCAGG 478

Qy    21 AsnValGlyValTrpAlaSerGlyValGluGlyValLysLysValAspIleAlaGlyTyrPro 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   479 AATGTGGGGGTCTGGGCCTCAGGGGAGGGGAAGAAAGGTGGGACATTGCGGGTATCTACCCC 538

Qy    41 ProValThrThrProPheThrAlaThrAlaGlnValValAspTyrGlyLysLeuGluGluAsn 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539 CCTGTGACCACCCCCTTCACTGCCACTGCAGAGGTGGACTATGGGAAACTGGAGGAGAAT 598

Qy    61 LeuHisGlyLeuGlyThrPheProPheArgGlyPheValValGlnGlySerAsnGlyGlu 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599 CTGCACAAACTGGGCACCTTCCCCTTCCGAGGCTTCGTGGTCCAGGGCAGCAATGGGGAG 658

Qy    81 PheProPheLeuThrSerSerGluArgLeuGluValSerArgArgValArgGlnAlaMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659 TTTCCTTTCCTGACCAGCAGTGAGCGGCTGGAGGTGTCGCGCAGGGTCCGCCAGGCCATG 718

Qy   101 ProLysAsnArgLeuLeuLeuLeuAlaGlySerGlyCysGluSerThrGlnAlaThrValGlu 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   719 CCCATGAACAGGCTCCTGCTACTGGCTGGCTCCGGATGCGAGTCCACTCAAGCCACAGTGGAG 778

Qy   121 MetThrValSerMetAlaGlnValAlaAspAlaAlaMetValValThrProCysTyr 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   779 ATGACCGTCAGCATGGCCCAGGTCGGGGGCTCGACCGCGGCCATGGTGGTGACCCCTTGCTAC 838

Qy   141 TyrArgGlyArgMetSerSerAlaAlaLeuLeuHisHisThrTyrThrLysValAlaAspLeu 160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   839 TATCGTGGCCGCATGAGCAGTGCGGCCCTCATTCACCATACACACGTTGCTGGATCTC 898

Qy   161 SerProIleProValLeuTyrSerValProAlaAsnThrGlyLeuAspLeuProVal 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   899 TCTCCAATCCCTGTGGTGTACAGTGTCCCAGCCAACACAGGGCTGGACCTGCCTGTG 958

Qy   181 AspAlaValAlaValThrLeuSerGlnHisProAsnIleValGlyMetLysAspSerGlyGly 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   959 GATGCAGTGGCTGTCACGCTTTCCCAGCACCCGAATATTGTGGGCATGAAAGACAGCGGTGGT 1018

Qy   201 AspValThrArgIleGluGlyIleLeuValHisGlyThrArgLysGlyGlnValValLeu 220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1019 GATGTGACCAGGATTGGGCTGATTGTTCAAGACCAGGAGGATTTTCAGGTGTTG 1078

Qy   221 AlaGlySerAlaGlyPheLeuMetAlaSerThrTyrAlaLeuGlyAlaAlaValGlyGlyValCys 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1079 GCTGGATCGGCTGGCTTCTGATGGCCAGCACCTATGCCCTTGGGGCTGCTGTGGGGGGTCTGC 1138

Qy   241 AlaLeuValLaAsnValLeuGlyValAlaGlnValCysGlnLeuGluGluArgGlyLeuCysGlyThrGly 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1139 GCCCTGGCCAATGTCCTCGGGGTCGCCCAGGTGTGCCAGCTGGAGGAGCGGGGC 1198

Qy   261 GlnTrpGluAspAlaGlnLysLeuGlnHisArgLeuIleGluLeuProAsnAlaAlaValThr 280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1199 CAATGGGAAGATGCCCAGAAACTGCAGCACCGCCTCATTGAGCCCAATGCCGCGGTGACC 1258

Qy   281 ArgArgPheGlyIleLeuProGlyLeuLysGlySerIleMetAspTrpPheGlyTyrTyrGlyGly 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1259 CGGCGCTTTGGGATCCCAGGGCTGAAGAAAATCATGGACTGGTTCGGCTACTATGGAGGC 1318

Qy   301 ProCysArgGlyAlaProLeuGlnGluLeuSerProAlaGluGluGlnGluValAlaLeuArgMetAsp 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1319 CCCTGCCGCGGCGCCCCCTTGCAGGAGCTGAGCCCCGCCGAGGAGCAGGAGGCACTGCGGATGGAT 1378

Qy   321 PheThrSerAsnGlyTrpLeu 327
         |||||||||||||||||||||
Db  1379 TTCACCAGCAACGGCTGGCTC 1399
```

RESULT 2
; US-10-511-270-1
; Sequence 1, Application US/10511270
; Publication No. US20050214858A1
; GENERAL INFORMATION:
; APPLICANT: NAKAJIMA, HIDENORI
; APPLICANT: OHKUBO, MITSURU
; APPLICANT: YOSHIMURA, SEJI
; APPLICANT: NISHIO, NOBUYA
; APPLICANT: NISHIO, KAORI
; TITLE OF INVENTION: NOVEL 35 KD PROTEIN
; FILE REFERENCE: 260617US0PCT
; CURRENT APPLICATION NUMBER: US/10/511,270
; CURRENT FILING DATE: 2004-10-20
; PRIOR APPLICATION NUMBER: PCT/JP03/05431
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: JP 2002-126107
; PRIOR FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 1
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-270-1

US-10-511-270-2 (1-327) x US-10-511-270-1 (1-1061)

Alignment Scores:

Pred. No.:	2.48e-189	
Score:	1659.00	Length: 1061
Percent Similarity:	99.4%	Matches: 326
Best Local Similarity:	99.4%	Conservative: 0
Query Match:	97.5%	Mismatches: 1
DB:		Indels: 2
		Gaps: 0

```
Qy     1 MetLeuGlyProGlnValTrpSerSerValArgGlnGlyLeuSerArgSerLeuSerArg  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8 ATGCTGGGTCCCCAAGTCTGGTCTTCGTGTAGGCAGGGGGCTCAAGCGGAGCTTGTCCAGG 67

Qy    21 AsnValGlyValTrpAlaSerGlyValGluGlyValLysLysValAspIleAlaGlyTyrPro 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    68 AATGTGGGGGTCTGGGCCTCAGGGGAGGGGAAGAAAGGTGGGACATTGCGGGTATCTACCCC 126

Qy    41 ProValThrThrProPheThrAlaThrAlaGlnValValAspTyr-GlyLysLeuGluGluAsn 60
         ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   127 CCTGTGACCACCCCCTTCACTGCCACTGCAGAGGTGGACTATGGGAAACTGGAGGAGGAGA 186

Qy    60 nLeuHisGlyLeuGlyThrPheProPheArgGlyPheValValGlnGlySerAsnGlyGl 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 TCTGCACAAACTGGGCACCTTCCCCTTCCGAGGCTTCGTGGTCCAGGGCAGCAATGGGGA 246
```

OM protein - nucleic search,

Run on: February 28, 2006, 11:51:51

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Search time 6074.23 Seconds
(without alignments)
3060.111 Million cell updates/sec

Title: US-10-511-270-2
Perfect score: 1702
Sequence: 1 MLGPQVWSSVRQGLSRSLSR..........ELSPAEEALRMDFTSNGWL 327

Scoring table: BLOSUM62
               Xgapop 10.0 ; Xgapext 0.5
               Ygapop 10.0 ; Ygapext 0.5
               Fgapop   6.0 ; Fgapext 7.0
               Delop   6.0 ; Delext   7.0

Searched: 5883141 seqs, 2842172565 residues

Total number of hits satisfying chosen parameters: 11766282

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abs/ABSWEB.spool/US10511270/runat_27022006_164246_19980/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss03h
-USER=US10511270 @CGN_1_5142 @runat_27022006_164246_19980 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
           1: gb_ba:*
           2: gb_in:*
           3: gb_env:*
           4: gb_om:*
           5: gb_ov:*
           6: gb_pat:*
           7: gb_ph:*
           8: gb_pr:*
           9: gb_ro:*
          10: gb_sts:*
          11: gb_by:*
          12: gb_un:*
          13: gb_vi:*
          14: gb_htg:*
          15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1702	100.0	2480	8	BC045550	Homo sapi
2	1641.5	96.4	1521	6	AR541908	Sequence
3	1516	89.1	1618	9	BC016430	Mus muscu
4	1441.5	84.7	265464	14	HSA312051	Homo sapi
5	1430	84.0	1336	6	AX458350	Sequence
6	1418.5	83.3	884	6	CQ722087	Sequence
7	1182.5	69.5	1148	5	BC087798	Xenopus
8	1172	68.9	1160	5	BC084421	Xenopus l
9	1155	67.9	1370	6	BC066708	Danio rer
10	1034	60.8	2581	6	AX833220	Sequence
11	1034	60.8	2581	8	AK094791	Homo sapi
12	791.5	46.5	1978	8	BC057821	Homo sapi
13	791.5	46.5	2012	8	BC011916	Homo sapi
14	755	44.4	933	5	CR407189	Gallus ga
15	738	43.4	505	5	BC011539	BST and e
16	738	43.4	505	5	BD111539	Homo sapi
17	738	43.4	505	6	AX972820	Sequence
18	605.5	35.6	157193	5	AJ355315	Human DNA
19	568	33.4	265464	14	AC157384	Bos tauru
20	564.5	33.2	149854	9	AC117225	Mus muscu
21	564.5	33.2	161150	9	AC133503	Mus muscu
22	564.5	33.2	223987	14	AC147800	Mus muscu
23	536.5	31.5	133470	9	AC131867	Mus muscu
24	536.5	31.5	222812	14	AC158177	Rattus no
25	394.5	23.2	210030	1	AC158177	Rattus no
26	382	22.4	16155	1	AY596297_01	Continuation (2 of
27	375.5	22.1	16155	6	AB004999	Halobacte
28	375.5	22.1	110000	1	AB017225_36	Continuation (37 o
29	375.5	22.1	161150	1	AB017334_36	Continuation (36 o
30	375.5	22.1	223987	1	AB017355_36	Continuation (37 o
31	375.5	22.1	291030	1	CP000001_36	Continuation (37 o
32	375.5	22.1	302481	1	AB017036	Bacillus
33	375	22.0	14300	8	AB017010	Bacillus
34	370	21.7	909	6	HSA304804	Homo sapi
35	365	21.4	1100000	6	AR347015	Sequence
36	365	21.4	1100000	1	AB017333_33	Continuation (34 o
37	361.5	21.2	1100000	1	CP000002_33	Continuation (34 o
38	361.5	21.2	1100000	1	AB017333_35	Continuation (36 o
39	356	20.9	5240	6	AR310393	Sequence
40	356	20.9	5240	15	AF056977	Penicilli
41	355.5	20.9	287765	1	AF056977	Penicilli
42	355	20.9	45055	14	AY714867	Unciture
43	353	20.7	11714	1	AB013350	Methanosa
44	352	20.7	4183	1	AF319635	Thermococ
45	347.5	20.4	1100000	1	AP008226_06	Continuation (7 of

ALIGNMENTS

RESULT 1
BC045550
LOCUS   BC045550   2480 bp   mRNA   linear   PRI 28-JUL-2005
DEFINITION Homo sapiens chromosome 10 open reading frame 65, mRNA (cDNA clone
            MGC:57219 IMAGE:5274464), complete cds.
ACCESSION   BC045550
VERSION   BC045550.1 GI:28278549
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 2480)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,

102(a)

```
Query Match:        100.0%      Gaps:          0
Best Local Similarity: 100.0%   Indels:        0
Percent Similarity:   100.0%    Mismatches:    0
                                Conservative:  0
Score:              1702.00     Matches:       327
Pred. No.:          1           Length:        2480
Alignment Scores:   4.34e-124
```

ORIGIN

FEATURES   Location/Qualifiers

source   1..2480
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:57219 IMAGE:5274464"
 /tissue_type="Brain, hippocampus"
 /clone_lib="NIH_MGC_95"
 /lab_host="DH10B"
 /note="Vector: pBluescriptR"

gene   1..2480
 /gene="C1orf65"
 /db_xref="GeneID:112817"

CDS   344..1327
 /gene="C1orf65"
 /codon_start=1
 /product="hypothetical protein LOC112817"
 /protein_id="AAH45550.1"
 /db_xref="GI:28278550"
 /db_xref="GeneID:112817"
 /translation="MLGPQVWSSVRQGLSRSLSRNVGWWASGEGKKVDIAGIYPPVTT
PFTATARVDYGKLEENLHKLGTPFPRGFVVQSSNGEFPPLTSSERLEVSRVRQAMPK
NRLLLACSGGCESTQATVEMTVSMAQVGADAAMVVTPCYYRGRMSSAALIHHYTKVADL
SPIPVVLVSVPANTGLDLPVDAVVILSQHPNIVGMKDSGGGDVTRIGLIVHKTKKQDFQ
VLAGSAGRLMASYALGAVGGVCALANVLGAQVCQLERLCCTGQWEDAQKLQHRLIEPN
AAVTRRGIPGLKKIMDWFGYGYGGPCRAPLQELSPAEREALRMDFTSNGWL"
```

```
Db 344  ATGCTGGGTCCCCAAGTCTGGTCGTCTTCTGTGAGGCAGGGGGGCTAAGCAGCAGAGCTTGTCCAGG 403
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  20  MetLeuGlyProGlnValTrpSerSerValArgGlnGlyLeuSerArgSerLeuSerArg 20

Db 404  AATGTGGGGTGGTGGGCCTCAGGGAGAGGAAAGAAGGTGGACATTGCGGGTATCTACCCC 463
Qy  40  AsnValGlyTrpTrpAlaSerGlyGluGlyLysLysValAspIleAlaGlyIleTyrPro 40

Db 464  CCTGTGACCACCCCCTTCACTGCCACTGCACGAGTGGACTATGGGAAACTGGAGGAGAAT 523
Qy  60  ProValThrThrProPheThrAlaThrAlaArgValAspTyrGlyLysLeuGluGluAsn 60

Db 524  CTGCACAAACTGGGGCACTTCCCCTTCCCCAGGGGCTTCGTGGTCCAGGGCTCCAATGGCGAG 583
Qy  80  LeuHisLysLeuGlyThrProPhePheArgGlyPheValValGlnSerSerAsnGlyGlu 80

Db 584  TTTCCTTTCCTGACCAGCAGTGAGCGCCTCGAGGTGGTCAGCCGTGTGCGCCAGGCCATG 643
Qy 100  PheProPheLeuThrSerSerGluArgLeuGluValSerArgValArgGlnAlaMet 100

Db 644  CCCAAGAACAGGCTCCTGCTAGCTGGCTGCAGGTGCGAGTCCACTCAAGCACAGTGGAG 703
Qy 120  ProLysAsnArgLeuLeuLeuAlaGlySerGlyCysGluSerThrGlnAlaThrValGlu 120

Db 704  ATGACCGTCAGCATGGCCCAGGTCGGGGCTGACGCGGCCATGGTGGTGACCCCTTGCTAC 763
Qy 140  MetThrValSerMetAlaGlnValGlyAlaAspAlaAlaMetValValThrProCysTyr 140

Db 764  TATCGTGGCCGCATGAGCAGTGCGGCGCTCATTCACCACTACACCAAGGTTGCTGGATCTC 823
Qy 160  TyrArgGlyArgMetSerSerAlaAlaLeuIleHisHisTyrThrLysValAlaAspLeu 160

Db 824  TCTCCAATCCCTGTGGTCCTGGTTAGCGTGTCCCAGCACACAGGGGCTGGACTCCTGTG 883
Qy 180  SerProIleProValValLeuValSerValProAlaAsnThrGlyLeuAspLeuProVal 180

Db 884  GATGCAGTGGTTGCACGCTTTTCGAGTGGCCAGCTATGCCTTGGGATGCGGGGCTGC 943
Qy 200  AspAlaValValIleLeuSerGlnHisProAsnIleValGlyMetLysAspSerGlyGly 200

Db 944  GATGTGACCAGGATTGGGCTGATTGTTCACAAGACCAAGAAGCAGGATTTTCAGGGTTG 1003
Qy 220  AspValThrArgIleGlyLeuIleValHisLysThrLysLysGlnAspPheGlnValLeu 220

Db 1004 GCTGGAAGTGCGGGCTTTCTTCGATGGCCGCCAGCTATGCCTTGGGGGTGCGGGGCCTGC 1063
Qy 240  AlaGlySerAlaGlyPheLeuMetAlaSerTyrAlaLeuGlyAlaValGlyGlyValCys 240

Db 1064 GCCCTGGCCAATGTCCTGGGGGCTCAGGTGTGCCAGCTGGAGCGACTGTGCTGCACGGGG 1123
Qy 260  AlaLeuAlaAsnValLeuGlyAlaGlnValCysGlnLeuGluArgLeuCysCysThrGly 260

Db 1124 CAATGGGAAGATGCCCAGAAAACTGCAGCACCGCCTCATTGAGCCCAAACCGTCGCGGTGACC 1183
Qy 280  GlnTrpGluAspAlaGlnLysLeuGlnHisArgLeuIleGluProAsnAlaAlaValThr 280

Db 1184 CGGCGCTTGGGATCCCAGGGCTGAGAGAAAATCATGGACTGGTTTGGCTACTATGGAGGC 1243
Qy 300  ArgArgGlyIleProGlyLeuLysLysIleMetAspTrpPheGlyTyrGlyTyrGlyGly 300

Db 1244 CCCTGCCGGGCACCCTTGCAGGAGCTGAGCCCTGCAGAGAGGGAGGCACTGCGCATGGAT 1303
Qy 320  ProCysArgAlaProLeuGlnGluLeuSerProAlaGluArgGluAlaLeuArgMetAsp 320

Db 1304 TTCACCAGCAACGGCTGGCTC 1324
Qy 321  PheThrSerAsnGlyTrpLeu 327
```

RESULT 2

```
LOCUS       AR541908               1521 bp    DNA    linear   PAT 08-OCT-2004
DEFINITION  Sequence 156 from patent US 6743619.
ACCESSION   AR541908
VERSION     AR541908.1  GI:53933988
```

102(a)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2005, 09:41:19 ; Search time 4837.81 Seconds
(without alignments)
10626.906 Million cell updates/sec

Title: US-10-511-270-1
Perfect score: 1061
Sequence: 1 gaagtctatgctgggtcccc..........gaacttgcagcctgaattcc 1061

Scoring table: IDENTITY NUC
Searched: Gapop 10.0 , Gapext 1.0
4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1025.4	96.6	2480	9	BC045550	Homo sapi
2	818.6	77.2	1442	9	HSA312051	Homo sapi
3	814.2	76.7	1336	9	AJ312051	Homo sapi
4	782.4	73.7	1521	6	AX458350	Sequence
5	761	71.7	884	6	AR541908	Sequence
6	722.8	68.1	1618	10	CQ722087	Sequence
7	445.6	42.0	2581	9	BC016430	Mus muscu
8	445.6	42.0	2581	9	AX833220	Sequence
9	422.6	39.8	1160	5	AK094791	Homo sapi
10	405	38.2	1370	5	BC084421	Xenopus l
11	351	33.1	2012	9	BC011916	Homo sapi
12	349.4	32.9	1978	9	BC057821	Homo sapi
13	329.4	31.0	933	5	BC066708	Danio rer
14	317.2	29.9	505	5	CR407189	Gallus ga
15	317.2	29.9	505	6	AR415986	Sequence
16	317.2	29.9	505	6	AX972820	Sequence
17	217	20.5	157193	6	BD111539	Sequence
c 18	189	17.8	14300	9	AL355315	Human DNA
c 19	137.6	13.0	149854	10	AC117225	Mus muscu
c 20	137.6	13.0	161150	10	AC133503	Mus muscu
c 21	137.6	13.0	223987	2	AC147800	Mus muscu
c 22	129.4	12.2	133470	10	AC131867	Rattus no
c 23	123.8	11.7	16155	1	AB004999	Halobacte
c 24	112.6	10.6	458	11	G80217	S208P6433PC
25	95.4	9.0	110000	1	AB017180_01	Continuation (2 of
26	89.2	8.4	10722	6	AB011168	Methanosa
27	85.6	8.1	346362	1	BX640425	Bordetell
28	85.6	8.1	348257	1	BX640425	Bordetell
29	81.2	7.7	11714	1	AB013350	Methanosa
30	79.6	7.5	879	6	CQ878055	Sequence
31	79.6	7.5	12967	1	AB004533	Pseudomon
32	78	7.4	876	6	CQ878057	Sequence
33	75	7.1	11371	1	AB012275	Xanthomon
34	74.4	7.0	5240	1	AR310393	Sequence
35	74.4	7.0	5240	8	AR310393	Sequence
36	73.6	6.9	349580	1	AF056977	Penicilli
37	71.2	6.7	10293	1	BX572601	Rhodopseu
38	70.6	6.7	301925	1	AB011807	Xanthomon
39	70.6	6.6	272101	1	AP005046	Streptomy
40	70.4	6.6	849	6	BD179647	Highly th
41	69	6.5	346879	1	AB017302	Thermus t
42	68.6	6.5	283100	1	BX572598	Rhodopseu
43	68.6	6.4	110000	1	AL939110	Streptomy
44	66.4	6.3	110000	1	AB017333_35	Continuation (41 o
45	66.4	6.3	110000	1	CP000002_35	Continuation (36 o

ALIGNMENTS

RESULT 1
BC045550
BC045550
LOCUS   BC045550   2480 bp   mRNA   linear   PRI 30-JUN-2004
DEFINITION Homo sapiens chromosome 10 open reading frame 65, mRNA (cDNA clone MGC:57219 IMAGE:5274464), complete cds.
ACCESSION   BC045550
VERSION   BC045550.1 GI:28278549
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2480)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED   12477932
REFERENCE   2 (bases 1 to 2480)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (31-JAN-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA